DOCUMENT RESUME

ED 125 297 FL 007 819

AUTHOR Ellis, Jeffrey
TITLE The Role of the Concept of Text in the Elaboration of Linguistic Data. York Papers in Linguistics, No. 6.
INSTITUTION York Univ. (England). Dept. of Language.
PUB DATE Mar 76
NOTE 18p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Descriptive Linguistics; *Haiku; *Linguistics; *Linguistic Theory; *Literary Analysis; Poetry; *Textual Criticism; Theoretical Criticism
IDENTIFIERS *Text Linguistics

ABSTRACT

Text linguistics, though firmly based on all precedent work in sentence linguistics, is not a mere extension of the methods of the latter to units above the sentence; it adds a new dimension to linguistics, which may affect our thinking about any aspect of linguistics. In this paper the concept of text is applied to the linguistic analysis of selected haiku. These poems are examined from three perspectives: (1) being "a" text, the demarcation of one text from another; (2) having the quality of being "text," the aspect or function of language that constitutes this; and (3) having the structure of a text. Peripheral topics, such as mixed texts, informant texts, and decontextualized texts, are also discussed. (DB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED125297

THE ROLE OF THE CONCEPT OF TEXT IN THE ELABORATION OF LINGUISTIC DATA*

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Jeffrey Ellis

(University of Edinburgh)

Text linguistics[1] is not a revolution in linguistics in the sense of being in any way a negation of previous linguistics; on the contrary it is firmly based on all that has been done in what I will call, following various writers, sentence linguistics. But it is not a mere extension of the methods of sentence linguistics to units above the sentence; it does add a new dimension to linguistics, which may affect our thinking about any aspect of linguistics.

One aspect, I suggest, is the nature of linguistic data. It would not be true to say that before Harris nobody thought textually or there was no textual aspect to the apprehension of linguistic data; but it is a question of a greater consciousness and of systematically developing the theory of linguistics to accomodate what had always been there in practice.

However, if I had not been specifically invited to contribute to this symposium and thereby challenged to make a detailed connection between some of my research interests and the subject of 'the nature of the data of linguistics' I would not have thought of making so explicit the role of the concept of text in the elaboration of linguistic data. I hope that my remarks will be taken in this spirit, as tentative and provisional, while the linguistic study of text, and indeed the philosophy of science of linguistics, are still in a formative stage.

FL007819

As to this latter, philosophy of science of linguistics, I shall have to take for granted much about the classification of the levels of abstraction involved in the observation, recording and processing of linguistic data, and about the levels or strata etc. of linguistic analysis itself and the branches of linguistic theory and associated disciplines, and the way in which the findings of one of these levels or branches are the data of another, with varying degrees of feedback.

To these levels of abstraction and other aspects of linguistic methodology is to be related the concept of text in various senses or aspects, of which I will distinguish three:

1. being *a* text, the demarcation of one text from another
2. having the quality of being *text*, the aspect or function of language that constitutes this
3. having the *structure* of a text.

To take these in turn:-

1. *Being a text* When we demarcate the flow of speech in its various manifestations into distinguishable occurrences of language, we do so partly because these occurrences have distinct situations, partly because of linguistic features. Consider for example the task of re-separating the little texts and bits

of text I have run together in the Appendix. Here of course, as often with written texts, the relevant situational factors are to be derived from the understanding of the words, as will appear from the examples discussed below.

Halliday (forthcoming: §3.2) makes the point that 'we should stress the essential indeterminacy of the concept of "a text"', as compared with such units as clauses or syllables. 'A text, in the normal course of events, is not something that has a beginning and an ending. The exchange of meanings is a continuous process that is involved in all human interaction: it is not unstructured, but it is seamless, and all that one can observe is a kind of periodicity in which peaks of texture alternate with troughs -- highly cohesive moments with moments of relatively little continuity. The discreteness of a literary text is untypical of texts as a whole.'

This formulation assumes that the normal use of language is conversation, and we shall be returning below to the question of the 'primacy' of spoken language, of dialogue, and of non-literary language. Meanwhile we note that *some* occurrences of language *can* be neatly demarcated into texts, and *all* occurrences, even conversations, can be subjected to *some* kind of demarcation, if we make the proviso that the demarcation of texts may be multi-layered.[2]

Take for example the first text in the Appendix, lines or paragraphs 1 and 2.

At this point I should say something generally about the examples in the Appendix. I had to find concise examples, not the easiest of tasks in the very field that concerns stretches of language at their longest, though it is true that as Halliday says (forthcoming: §4.1) 'Many things about language can be learnt only from the study of very long texts. But there is much to be found out also from little texts; not only texts in the conventional forms of lyric poetry, proverbs and the like, but also brief transactions, casual encounters, and all kinds of verbal micro-operations.'

I was prompted to exploit the haiku (and senryuu), those possibly shortest of non-oral literary texts, for some of my examples by what is said about an English haiku[3] by Widdowson (1975:37-8), drawing on Alex Rodger. I will return to this below.

The lay-out of the examples, without punctuation or line spacing, is designed to give a sample, albeit artificial and arbitrary, of a continuum or flow of language without what is specific either to the usual written presentation, punctuation within the text or separating of the texts, or to a given spoken realization of them, intonation. I *have* kept the sources' line-division of the verse, with for the Japanese word-for-word, sometimes morpheme-for-morpheme, glossing, and the paragraph-division of the Thurber prose and speaker-division of the dialogue 11.24-27. A key is provided at the end.

To take the Thurber example (11.1-2), then, as a text in itself -- it represents in fact the beginning of a longer text, a news report Thurber says he wrote after his editor had told him to write shorter lead sentences to his crime stories. Its linguistic features (which we shall be coming to under informational structuring) make sense once one sees it as a text in this situation or as a constituent text embedded as reported writing in the larger text of the account of Thurber's exchange with his editor.

2. *The quality that language has as text* (This of course is interrelated with 1., its properties as *a* text demarcated from others.)

According to Halliday (forthcoming: §2.1) while 'a text is the product of meanings of all four kinds -- experiential, logical and interpersonal, as well as textual', the *textual* component of the lexico-grammar has the specific function 'of making the difference between language in the abstract and language in use', and in English at least consists of 'I) The structure-generating systems....(i) thematic systems and (ii) information systems, see Halliday 1968. II) The cohesive relations....(i) referential, (ii) substitutive-elliptical, (iii) conjunctive and (iv) lexical'[4] -- for a partial re-ordering of i, ii and iv see Ellis (forthcoming a and b).

As an example, let us consider the twenty words quoted by Widdowson (1975:37) which as I have said formed the starting-point of the bulk of the Appendix, of which they are 11.9-11. I say 'twenty words' and not 'text', because Widdowson denies that they form a text in the full sense I give the term (n.1), combining his 'text' and 'discourse'. He says (38): '...in spite of the fact that they do not form a text (being two noun phrases not syntactically linked as a sentence) they do nevertheless constitute a discourse'.

One textual feature of these two noun phrases plus the title phrase that we might point to, under Halliday's II) cohesive, (i) referential, is the phora of the deictics in them. For classification and terminology of phora see Ellis (1971) and Martin (forthcoming). The *the* in 'the metro' is homophoric, in 'the crowd' anaphoric to 'the metro' in the title or if we think of the title as meta-textual ecphoric to the crowd in the metro given as part of the poet's situation of reference, in 'the apparition' cataphoric or esphoric to 'of these faces', the *these* in 'these faces' is ecphoric.[5] But all the deictics in this phrase, together with one of the two in the title phrase, are definite, whereas there are no definite deictics in the final phrase and one indefinite one, so that part of the linguistic constituency of the whole is a contrast in phora. I will return to this under text-structure (3.) in considering the semantic structure of the text.

In introducing this example Widdowson, saying 'Consider, for example, this well-known *haiku* by Ezra Pound:', makes no mention of the structure of haikus in general -- whether formal-linguistic, semantic, rhetorical, poetic (including metrical and, to use Wexler's (1964) term, grammetrical), or discourse structure of any kind. In fact, as other haikus in the Appendix show, it is part of the genre structure (Halliday's 'generic structure' -- 'the form that a text has as a property of its genre') of most[6] haikus in the original Japanese to end in a noun or what Bownas and Thwaite term an 'emotional ejaculation' (1964:lxvii), i.e. a sentence-modifier or an interjection. An exception is 1.5 in the original, ending the haiku in a verb, if indeed it *is* a haiku and not a senryuu, senryuus differing inter alia in this feature.

On the other hand, not so many haikus have the grammatical structure with no verb anywhere that Widdowson describes as 'the simple juxtaposition of the two phrases' or (38) 'two noun phrases not syntactically linked as a sentence', and indeed the English of 11.14-16 has this absence of finite verb where the original Japanese has the finite verb in 1.15.

But both linguistic features (linguistic, hence textual in Widdowson's sense) of genre or sub-genre help us to see that this text (or 'discourse')

begins in 1.10, or 1.9 if we include the title included by Widdowson, and ends in 1.11, and does not fall apart between 11.10 and 11.

3. *Text structure* This is an aspect of 2., being text, but is different from the below-sentence structure of theme and information included there, and is also to be distinguished from the particular case of genre structure there. In fact, the question how far texts as wholes have a *linguistic* structure as distinct from a rhetorical structure, a poetic structure, and so on, and what the relation is between logical structure of a text, both in the non-linguistic sense and in Halliday's sense of the logical function of language, and a textual structure in the sense of his textual function of language, is a complex question and one of the controversial questions in text-linguistics. For some discussion of this see Ellis (forthcoming a and b).

If a text has structure in the linguistic sense in which a sentence has grammatical structure, then one might expect such structure to have a rank-scale with units and systems of classes of units, distinct from, though they may correlate with, the higher units of phonology (like the so-called 'paraphone'), and of graphic analysis (like the paragraph, chapter and so on), and indeed of grammar in the sentence-linguistics sense. This scale may, and in fact seems to be found to, differ for different genres, and this fact may reinforce suspicions that it is not a purely linguistic structure that is in question. But it *is* found, in genres so far investigated and reported on, to be similar in its lower units and to diverge in the higher ones, as well as in classes of units within each genre's repertory, which is not surprising when one considers the relation between simple speech acts and texts of lesser or greater degrees of higher complexity, including the multi-layered demarcation of texts already referred to (n.2). Thus Sinclair and Coulthard (1975) for spoken interaction between teacher and class, and Jones for written scientific exposition, both have as lowest units of discourse analysis act and move[7], but then as higher units Sinclair and Coulthard have exchange, transaction (or sequence), lesson (or interaction)[8], while Jones has stage and plan.[9]

Let us look briefly at the possible kinds of textual structure of the genre haiku as regards the English haiku 11.9-11 (cf.n.3) and the others in the Appendix, noting first the point of sentence-grammar structure in 11.9-11 that 1.10 is a nominalization of either 'these faces in the crowd appear' or 'these faces appear in the crowd'.[10]

One structure of the content of a haiku, according to Bownas and Thwaite (1964:lxvi-lxvii), in the sub-genre initiated by Bashoo, is one of three parts, exemplified by 11.17-19, where the first line represents 'the unchanging', the second line 'the momentary', and the last line '(the splash,) the point of intersection of the two', 'the intuitive flash of Zen', as they say. We might call this an *event* haiku.

Another obvious structure of content is one of two parts, exemplified by the other haikus. The two parts are in a quasi-predicative relation of topic and comment, or the relation in a metaphor between tenor and vehicle, or that between a riddle and its answer; and the sequence may be either topic -- comment as in 11.12-13 with comment indefinite '*a* butterfly' (or generic if sequence riddle--answer) or vehicle--tenor as in 11.3-5 (and comment--topic 11.14-16 if translated '*the* dragonfly' -- in Henderson's translation 'dragonfly' without deictic), or ambiguous as 11.6-8 (or 11.14-16 in Henderson's translation). For the dual possibility of sequence we

may compare the contrasting sequence of the simile or metaphor form familiar in English, e.g. '(the) boys are (run like) lightning', and that of Western Apache reported by Basso (1975), 'lightning is a boy' or 'carrion beetles are white men'. We might call these *static image* haikus. Some have a perception word, in original or translation, like 1.3 'feel', 12b 'mireba', Bownas and Thwaite 'Fallen flower I see', Miner 'returning I see A fluttering butterfly', or 1.15, Japanese student's translation, 'I saw reflected in its eyes', as indeed the event haiku 1.19 has the final perception noun 'sound'.

Pound's haiku 11.9-11 clearly belongs, at one level of interpretation, to the topic--comment sequence type, with the deictics as we have said (above, 2.) making the topic definite, the comment indefinite. At the same time one might conceivably interpret it as having present, though not neatly distributed between three lines of the verse, the three elements of an event haiku: the unchanging, namely the classical Japan of the traditional haiku, represented by the petals on the bough,[11] the momentary, the modern world, represented by the title and 'these faces in the crowd', and the intersection, represented by the perception-word (though nominalized) 'apparition' and by the linguistic mixture in register (see below). But the outer form at least is static image haiku, as we might expect from Pound's well-known general preoccupation with Chinese-character images.[12]

How far these structures of content could be shown to have corresponding features of language, as some of Sinclair and Coulthard's 'discourse classes' or Jones's 'rhetorical strategies' (with underlying 'conceptual structures') have what Jones calls 'grammatical realizations' corresponding, is a difficult question, impossible to answer on the little evidence to hand, but one that would have to be resolved before finally relating haiku-structure to a general text rank framework of act, move or whatever appropriate units.[13]

If we now begin to relate these aspects of the text concept to the levels of abstraction in the treatment of linguistic data, more explicitly than in the examples of the aspects of text so far, an important point about the status of the linguistic unit *text* in the organization of data concerns the abstraction of *registers* (varieties according to use, 'diatypic varieties') from the primary-data *continuum* of language-variety.

For the general problem of the recalcitrance of the continuum of language-variation to establishment of norms, local, social or 'diatypic', see for example Le Page (1969, 1973, 1975a and b), De Camp, Bailey, Hymes, Labov in Hymes (ed. 1971), Labov (1970), Sankoff (1972, 1974), Bickerton (1975), Ellis and Ure (forthcoming: §§3.15, 3.222).

Now whether or not one agrees with Halliday (forthcoming: §3.2) that text is a *semantic* unit (determined by multidimensional properties rather than linearly structured) -- and that register (§4.2) is a *semantic* system whose entry-point is this unit --, it is clear, I think, that the syntagmatic cut into texts brings about the paradigmatic cut into registers, inasmuch as it is when we compare the range of language-variation within one text with that in others, with similar or different situations of use, if we take sufficiently large and numerous samples, with sufficiently distinguished situational factors, that we can discern broad patterns within the overall welter of variation. It may be arguable that such samples have not yet been worked on enough to prove finally the validity of register in

any one language, but not that register has been proved to be *invalid*. It is also true that if text-demarcation as a datum brings about the analyst's distinction of register or 'genre structure', the latter too as we have seen, when generalized from the data of many demarcated texts, help to bring about the analyst's demarcation of a text in particular cases. Such is the kind of feedback between stages of data-handling already mentioned.

Looking at the Appendix, and leaving aside the mode distinction of written and spoken to which we shall be coming below, we can see in the texts certain groupings of genre, formality and field, most generally divisions that we may call provisionally 'journalism register' (11.1-2), 'haiku register' (11.3-19) (meaning register or registers used in English translation of haikus, not necessarily peculiar to it), and 'conversation register' (11.23-29) from three local register-systems, Scotland, Ghana and Wales, and something in between in 11.20-22. Within the 'haiku register' 1.9 differs in lexical universe of discourse from what we expect from classical haikus, and it is arguable that 11.9-11 altogether differ from original Japanese haiku genre structure in not containing a word identifying one of the four seasons (*kigo*).[14] But of course with such short texts, and so small a sample from the body of haikus in Japanese and in English of translation or derivative composition, nothing remotely approaching the statistical analysis that a full treatment of register demands is possible.

Registers differ not only in clear-cutness of text-demarcation (1.), but also in (2.) the incidence of linguistic features from the textual component of the language generally, and specifically (3.) text structure. The most fundamental difference of this kind is that between spoken and written language,[15] and this is also the most fundamental difference in the nature of linguistic data: spoken and written data call for differential application of the text concept already at the lowest level, as we have seen in considering text-demarcation; at the recording and storage stage spoken data may become 'written' by being transcribed or remain only spoken on tape: Sinclair (*et al.* 1972: §1.6) discusses briefly kinds of purposes in transcribing spoken texts in toto, the problem of paralinguistic features, and other demands on editorial responsibility which may in fact anticipate the 'processing' levels (an example in the Appendix is in 1.27, '...ɛ...ɛyɛ', where I have kept points of suspension in the transcript indicating hesitations in which ɛ- ('it') is repeated); at the 'processing' levels spoken data do become 'written', e.g. in broad transcription 1.23, or transcribed in traditional orthography 11.24-29, but still differ from the data of written language.

An example of punctuation differentiating written data is concealed in 11.12-13 of the Japanese, which has two interpretations that would be punctuated differently if punctuated at all, represented by Miner's translations as

> To the branch of falling flowers
> Seems to return (a flower):

and

> To the branch a fallen flower
> Returns; when I look,

(in the first interpretation the first two words in Japanese romanization are

hyphenated, though not of course in characters -- nor if they were to be written in kana).

This word-play (*kake-kotoba*, 'pivot-word') is a possible feature of some Japanese poetic genre structures, in which a single word creates a double text.

At the level of linguistic, and sociolinguistic (or psycholinguistic), conclusions, the text concept is crucial in further differentiating of spoken and written.

A central case of this difference between spoken and written is informational structuring, Halliday's thematic and information systems, the latter, information structure, in Halliday's sense appearing usually only in spoken, since it is a matter of the aspect of intonation not shown in writing as the demarcation of information *units* (tonality) may be by punctuation (Halliday, 1967:200) -- an exception is when italics (or particularly underlining in intimate letter register) are used for marked tonicity. So in written English theme and associated systems (see Halliday, 1967/8) have to be used in ways additional to their spoken use, replacing spoken use of intonational structuring of given and new. (Cf. Firbas 1972).

The ongoing work of H.M.P.Davies (e.g. 1975) is concerned with the relation between the spoken and the written in this respect, with particular reference to English in schools and the problem of writing to be read (aloud) intelligibly (one aspect of Abercrombie's 'spoken prose', see Abercrombie, 1965). Davies (1975) quotes Halliday's[16] example of a passage (spoken commentary) that is not a text, not (just) in Widdowson's sense, but because the relevant structures are not appropriate to their intersentential environment (Davies: 'fundamentally unreadable, not just badly written ... virtually impossible to find an intonation to use when reading it aloud'): e.g. (see n.16) *Now comes the president here. It's the window he's stepping through to wave to the crowd. On his victory his opponent congratulates him. What they are shaking now is hands. A speech is going to be made by him...*

In the example from written language in the Appendix 1.2, if the spoken realization is difficult it is not for the same reason as the Halliday example just quoted. Given the preceding paragraph, *Dead.*, only such a structure is possible for the necessary information presentation; what is unusual is to front-shift (even to thematize within the clause -- *'Dead is what...'*) an element like this from a sentence of this complexity.

We have been considering the application of the concept of text to data both of written and of spoken language, as being equal though different. But it might be objected that whereas a text is typically written (and premeditated), literary and by one hand, language is typically spoken (and spontaneous) non-literary dialogue. (As Halliday (forthcoming: §3.4) puts it, 'we are led naturally to consider spontaneous conversation, as being the most accessible to interpretation; and to draw a rather clear line between this and other, less immediately contextualizable acts of meaning such as a poem or a prose narrative.' On primacy of dialogue cf. Shukman (1975) on Voloshinov's view: '"Verbal interaction is the basic reality of language": dialogue should be the central topic of linguistics' ('Voloshinov examines in close detail the phenomenon of reported speech in liter-

ature, ways by which authorial and character speech are combined and differentiated.'); and Ellis and Ure (forthcoming: §2.11).)

It is true that language originated in the spoken form. More important than the relative antiquity of spoken and written, as Lyons says (1968: 39), is 'the fact that all systems of writing are demonstrably based upon units of spoken language', though this argument needs a little refining as regards the earliest type of writing, Lyons' 'ideographic' (better termed, especially for Lyons' argument, *logographic*), which is based on the morpheme (rather than the word, as Lyons says), the morpheme (or indeed the word as a grammatical unit) being, unlike syllable or phoneme, as much a unit of the written language as the spoken.[17]

(On the other hand, not only have the spoken and the written a development of their own, as Lyons acknowledges (1968:39-42, 63), but what become common developments may originate as innovations in either. The study of diglossia, historical and contemporary, and its resolution is instructive in this regard, cf. the references in Ellis and Ure (forthcoming).)

It is also true that the term *text* was originally, and still often is, applied to written language only. Even professional linguists may tend to the term *utterance* instead for spoken language. As we have seen, *utterance* is variously defined; in its more usual use (n.2) it refers to a constituent of spoken texts, especially dialogues. The term *discourses* might be proposed for these, but again *discourse* has various senses.

But the fact seems to be that from a descriptive point of view it is best to have a general term for language of any kind as organized by the textual function, and within this framework to look at the particular effects of the senses in which it is true that spoken, dialogue and non-literary are 'prior to' written, monologue and literature.[18] Indeed, so far has *text* as a technical term been removed from its original limitations of application that Halliday (forthcoming: §3.4) can say 'It is natural to conceive of text first and foremost as conversation: as the spontaneous interchange of meanings in ordinary, everyday, interaction.'

Moreover he says (§4.3) 'one of the effects of a sociosemiotic approach is to suggest that *all* language is literature, in this sense; it is only when we realize that the same things are true of the spontaneous verbal interaction of ordinary everyday life (and nothing demonstrates this more clearly than Harvey Sacks' brilliant exegesis of conversational texts, which is in the best traditions of literary interpretation) that we begin to understand how language functions in society -- and how this, in turn, has moulded and determined the linguistic system.'

Elsewhere, in §4.3, he relates literature to conversation, and incidentally fits instances of realization of literary text into larger, more transient, spoken text. Above I called haikus non-oral literature to distinguish them from such short texts as proverbs or riddles. But Halliday says 'Oral verse forms such as ballads, lyrics, and epigrammatic and allusive couplets figure in many cultures as modes of competing...',[19] and Bownas and Thwaite (1964:xxxvii-xxxix) describes the integration of poetry into Japanese life up to the present day in haiku clubs and so on.

I will conclude with some questions of 'text' peripheral to the central concept.

The *mixed* text problem as described by Hill (1958:450-452), 'The Analysis of Mixed Texts', is a question of local dialects, and specifically of koine like English in Scotland and dialect like Scots; but we can generalize it to any kind of language or language-variety, and it becomes the limiting case in the problem of language-variation and variety-continuum referred to in discussing text-demarcation and register above. It might be said that all texts are to some extent mixed, but if so there are degrees of this, which indeed the statistical approach to register incorporates; and the methodological argument needs to assume that we have the data of relatively unmixed texts yielding our varieties before turning to really mixed texts.

The Appendix exemplifies mixture in texts in various ways. If we were to take it as one continuous 'text' it would be mixed:-

a. as writing: conventional orthographies and IPA, following from b.
b. written language and 'spoken' recorded in 'writing'
c. languages (English, Scots, Akan, Welsh, Japanese)
d. language-varieties, including registers as well as local varieties of English.

Most of this is accounted for by its being not one real text but an assemblage for exemplificatory purposes of several real texts, but within two (at least) of these there is still mixing:-

in 1.23, of two closely related tongues, English and Scots -- in this particular specimen Hill sees only one Scots 'dialectism' (the ending of 'dʒagət) distinguishing it from the English of Scotland

in 11.24-27, of two quite unrelated languages, English and Akan, and here the register-range to which the register of the text belongs falls into two (or more) distinct languages, in a way the Scottish one does not, most registers in it being unmixed Akan (or other Ghanaian language) or unmixed English (of a Ghanaian variety). This special mixed register variety can, I believe, be demonstrated to belong, institutionally and, in some senses at least, formally, to the Akan half of the register-range. See Ure (1974a).

L.28 is English, but such English as only occurs in a home environment where the school medium of instruction in the subject (here arithmetic) is Welsh. L.29 is the corresponding Welsh in the reverse situation of the previous generation.

It is also possible that the *register* of each of another two of the texts might be found to be in some sense mixed. That of the English of ll. 20-22 is intermediate between haiku-translation register and conversational, and there might be enough body of similar text for the analyst to distinguish a senryuu-translation register or set of registers. That of 11.9-11 is English haiku register with some possible deviance.

There are two other types of data in some sense going under the name of 'text' that should be discussed in a full account of the text concept.

'Informant texts' are central to the data of linguistics, but I will

mention first 'decontextualized texts' because of their theoretical bearing on informant texts, although in general they are not primary data.

A decontextualized text is usually an exemplification of language not observed in use but devised for the purpose, and the language intended to be exemplified is usually a user-variety, and not a specific register. As Widdowson (1978: §6.2.3) says, a sentence as exemplification of an abstract linguistic element[20] and the representation of a potential utterance (locution) may look alike on the page. This is more than decontextualized, we might call it a detextualized sentence. Candidates in the Appendix are 11.23 and 28-29, though for practical purposes it may be regarded as an academic question whether Hill or Thomas actually recorded one specific occurrence of these eminently possible utterances. But others might not be so eminently possible, and one could theoretically construct a cline of decontextualizedness attempting to assess likelihood of sentences' occurring in use (in given registers), both from the textual and from other linguistic points of view.[21]

Informant texts are broadly of two kinds (again a cline between is possible). Traditionally they are a sort of decontextualized texts, with investigator and informant cooperating somehow in producing them, a frame of some sort eliciting a linguistic form or a suggested form eliciting acceptance or rejection. Even here either party might be said to be dealing in contexts or situations as well as forms. (To take 11.23 and 28-29 again, one can more easily imagine *elicitation* of 1.23 than of 11.28-29.) But increasingly use is made in linguistic, especially psycholinguistic and sociolinguistic, research of informant data in a wider sense, and one feature of this symposium is some attention to method in such fields. The question here is to what extent such data qualify to be termed texts. In the case of the first kind of informant data, the product of the informant process is bits of text at least in the limited sense of being language, whether or not they are also Widdowson's 'discourse'. In the case of the second kind, what the analyst gets out of his informants may be text in this sense, and is more likely in the conditions of this kind of informant work to be also 'discourse'; but it may alternatively not be language as linguistic data at all but such psycholinguistic data as evidence of attitudes, which may be expressed in a language *of* description (rather than *under* description), or such sociolinguistic data as language-diaries, written in a language possibly different from that of the event recorded, though it is possible as part of the processing stages to include treatment of this 'meta-text' as language *under* description, for special purposes of research-category-identification, as in Ure (1974b, and also 1973).[22]

APPENDIX

1 dead
2 that was what the man was the police found in an areaway last night
3 what piercing cold I feel
4 my dead wife's comb in our bedroom
5 under my heel
6 scattered petals lie
7 on the rice seedling waters
8 stars in the moonlit sky

9 in a station of the metro
10 the apparition of these faces in the crowd
11 petals on a wet black bough
12 the fallen blossom flies back to its branch
13 a butterfly
14 far off mountain peaks
15 reflected in its eyes
16 the dragonfly
17 an old pond
18 a frog jumps in
19 sound of water
20 the tram car full
21 stop shoving they shout
22 and go on shoving
23 av ˈdʒagət ma ˈpɪŋke
24 misuro every inch of that road
25 yɛse two a die at once baako wui on the spot na baako nso yɛse owuu Nkawkaw
26 hyiee
27 ɛnna driver no nso yɛse ne condition...ɛ...ɛyɛ very serious
28 have you got any rhifyddeg tonight
29 oes arithmetic gen ti heno

Original of 11.3-8, 12-19:-

3 mi ni shimu ya
body at/to pierce '!/?'

4 boosai no kushi
late wife 's comb

5 neya ni fumu
bedroom step on

6 sakura chiru
cherry be shed

7 nawashiro mizu ya
rice-bed water

8 hoshi zukiyo
star moon night

12a rakka eda ni
fall blossom branch

12b kaeru to mireba
return *conjunction* look if

13 kochoo kana
butterfly '!'

14 tooyama no
far mountain *subject*

15 medama ni utsuru
eye gem be copied

16 tombo kana
dragonfly

17 furu ike ya
old pond

18 kawazu tobikomu
frog fly insert

19 mizu no oto
water 's sound

1-2 Thurber
3-5, 6-8 Yosa Buson, trsl. H.G.Henderson
9-11 Pound
12-13 Arakida Moritake, Pound's version
14-16 Kobayashi Issa, trsl. Bownas and Thwaite
17-19 Matsuo Bashoo, trsl. Bownas and Thwaite
20-22 modern senryuu, trsl. Bownas and Thwaite
23 Hill (1958:450)
24-27 Forson (1968)
28-29 Dr. C.Thomas (personal communication)

NOTES

* My thanks are due: for comments at all stages, to Miss J.N.Ure; for comments on the post-delivery draft, to A.Rodger; and for contributions to the discussion of the paper when delivered, to Mrs S.Y.Killingley, G.Sampson, and O.Uren.

1. By text, I should say for those familiar with Widdowson's terminology, I, with at least some other text-linguists, mean both what he calls 'text', the linguistic form of text, and 'discourse', text in relation to situation, meaning-context, etc.

On text linguistics in general, and diverse views on how the theory should be developed, see Dressler 1972, etc.

One crucial question, straddling text linguistics and sentence linguistics, is the status of the category *sentence* itself. This paper assumes the sentence: grammatically, as a structure composed of clauses (particular case of clause-complex); textually, as the minimum unit of language in use; orthographically, as demarcated by stops. Cf. fn.10.

2. As Harris says, quoted by Lyons (1968:172), 'Many utterances are composed of parts which are linguistically equivalent to whole utterances occurring elsewhere', and the relation between Harris's utterance, his definition of which includes its being a stretch of talk by one person, on the one hand, and a text which may involve more than one speaker, on the other hand, is one example of multi-layering.

Cf. Ure (1973, 1974b) on degrees of complexity in speech events.

At the same time, we have to distinguish smaller texts making up the layers of a larger text, e.g. hymns in a church service, from the embedding of a text as reported speech within the sentence-grammar of another text, as in Appendix, 1.21.

3. G.Sampson questions whether the term *haiku* is meaningfully applicable to any text not in Japanese. In so applying it I am following the usage of Pound's critics. (Note that Widdowson, for example, italicises the term.) If a haiku is defined metrically, not only do not all English translations or imitations of Japanese haikus have seventeen (5+7+5) syllables but an English syllable is in any case a unit in a different phonological, or phonologico-orthographical, system from the Japanese (where for example Appendix, 1.4 is seven or 1.16 five syllables). But it may be defined by other kinds of textual structure, on which see below.

4. He continues: 'see Gutwinski (in press), Halliday and Hasan (in press). There is no implication here that these are universal features; they may be, or they may not. But the systems in each network, and the way the systems are realized, are specific to the language in question.'

5. Esphoric (or eisphoric): a subdivision of 'cataphoric', pointing 'forward', namely pointing (not forward in all languages) into its own nominal phrase (to modifier or qualifier), not further in the text ('text-cataphoric', e.g. *its* in Appendix, 1.15).

The *these* in 'these faces' might seem to have the same phora as in conversational 'there was this crowd' (felt by some as obtrusive for its apparent pseudo-anaphora), but has arguably rather (depending partly on whether 'in the crowd' is qualifier or adjunct, cf. under 3.) the poignant immediacy of what might be termed poet's proximate deixis (as distinct from the distancing deixis, with possible nostalgic nuance, of **those faces*).

6. Bownas and Thwaite (1964:lxvii): 'the seventeen syllables should ideally -- and nearly always did -- end in a noun or an emotional ejaculation'.

7. for Sinclair and Coulthard correlating with grammatical clause and sentence respectively, and having classes of act such as *elicitation, prompt, reply,* of move such as *opening, answering, follow-up*; and in Jones's examples correlating with graphic sentence and more than one sentence respectively, and having classes of act like *ascription,* of move like *problem deduction.*

8. correlating with 'non-linguistic organization', transaction with topic, lesson with period, and exchange having primary classes *boundary, teaching.*

9. stage correlating in his examples with more than one paragraph, and having classes *problem identification, solution presentation.*

10. Note that in the former 'petals...' could conceivably be intensive complement to 'appear' but then with 'apparition' would have to be 'as petals...'. As it is, the two lines seem to be not one clause

but one sentence (cf. fn.1); it is of interest that in his discussion of the origin of the poem (1916:100-3) Pound himself called it a sentence ('...six months later I made a poem half the length; a year later I made the following sentence:'), no doubt in an orthographical sense -- I am indebted to A.Rodger for this reference.

11. although within Japanese culture petals or blossoms are a symbol of transience -- as in the medieval Iroha (the syllabary repertory of forty-seven kana making up a 'quick brown fox' text of four couplets of Buddhist verse): 'Iro wa nioedo Chirinuru wo -- Waga yo tare zo Tsune naran?...', 'Though gay in hue, they flutter down, alas! Who then, in this world of ours, may continue forever?...' (Chamberlain's translation).

12. although Pound himself thought of the pictures he saw in Chinese characters as dynamic, see e.g. Dembo (1963). (Cf. fn.17.)

13. For example, we might analyse a static image haiku as one move of either two acts or one, and an event haiku as one move of three acts; but if instead we re-assimilated the event haiku at one level of content to static image haikus by treating the splash as comment, we should have to analyse the whole as two moves, the first having two acts, making static image haiku structure two moves.

    It may be felt that haikus are not long enough to bear structural analysis of this order, echoing Brower and Miner's 'are the Japanese poetic forms long enough to be considered poems; and, if they are poems, how significant are they for Western readers?'

14. unless 'petals' would count. But part of the way in which the original nature of a haiku is reversed is that, as the title indicates, the scene set is in a place without seasons, and 'petals' belongs not to the topic but to the comment.

15. Indeed Halliday (forthcoming: §4.2) directly relates the textual function (or functional component of the semantics) and the register dimension of mode (as he does the interpersonal function and tenor, and the experiential function and field).

16. The example given by Davies is quoted from a lecture given by Halliday in 1969. A better version for our purposes appears in Halliday (forthcoming: §3.1) and is quoted here in part.

17. A relevant datum to set against Lyons' qualification (40) of priority of speech that 'Written and spoken Chinese are even more independent of one another' (than French, 41), and especially as regards the classical language that he does not specify, is the evidence that the writing of classical Chinese poetry is merely a coding of spoken poetry, and Pound's and others' attribution of poetic value to the visual constituency of the characters is fallacious. See Liu (1962), Dembo (1963), Killingley (unpublished).

    Another datum is that in the writing of Japanese, as distinct from Chinese, most characters have several readings, but in a Japanese poem a character encodes one single reading in the given text. (The ambiguity of *kake-kotoba* (see above) is quite another matter, being in the words of the spoken poem itself.) I am indebted to Mrs S.Y.Killingley

for raising the question of possible graphic factors in the translation of Japanese poetry.

18. Halliday (forthcoming: §3.4): 'It is perhaps useful in text studies... to bring out those aspects of the semiotic act that are common to all ...the very general concept of a text as an exchange of meanings...'

19. An example of collaborative emulation is the haikai by several hands, Bownas and Thwaite (1964:124-127).

20. which in Widdowson's terms has signification but not value, implication of utterance.

21. A study relating this to observed cases of decontextualized text in linguistic expositions would constitute a special case of 'register-description (and comparison)', of the language-variety of linguistic-exposition exemplification.

22. Without being able to devote space to describing an 'informant test' of yet another sort which contributed to the making of this paper, I should like to express my thanks to my poetic-textual informants Miss E.Araki and Miss M.Nakano.

REFERENCES

Abercrombie, D. (1965) 'Conversation and Spoken Prose', *Studies in Phonetics and Linguistics*. Oxford U.P.

Basso, K. (1975) '"Wise Words" of the Western Apache: Metaphor and Transformational Semantic Theory', Wenner-Gren Symposium on Semiotics of Culture and Language.

Bickerton, D. (1975) *Dynamics of a Creole System*. Cambridge, C.U.P.

Bownas, G. and Thwaite, A. (1964) *The Penguin Book of Japanese Verse*. London, Penguin Books.

Davies, H.M.P. (1975) 'A Way in to Language Study', DES Summer School for English Teachers, Oxford.

Dembo, L.S. (1963) *The Confucian Odes of Ezra Pound: a critical appraisal*. London.

Dressler, W. (1972) *Einführung in die Textlinguistik*. Tübingen.

Ellis, J. (1971) 'The Definite Article in Translation between English and Twi', in M.Houis (ed), *Actes du huitième Congrès de la Société Linguistique de l'Afrique Occidentale, Abidjan, 1969*. Abidjan.

(forthcoming a) 'The Logical and Textual Functions'.

(forthcoming b) 'Textual Meaning in an Akan Folktale and its English Translation'.

Ellis, J. and J.N.Ure (forthcoming) 'The Contrastive Analysis of Language Registers', in W.Winter and W.Nemser eds, *Trends in Contrastive Linguistics*.

Firbas, J. (1972) 'On the Interplay of Prosodic and Non-Prosodic Means of Functional Sentence Perspective', in V.Fried (ed), *The Prague School of Linguistics and Language Teaching*. Oxford U.P.

Forson, B. (1968) *Description of Language Situations Classified in a Corpus Representing a Bilingual Register Range (Akan Varieties and English).* University of Ghana, unpubl. MA thesis.

Gutwinski, W. (in press) *Cohesion in Literary Texts: a study of some grammatical and lexical features of English discourse.*

Halliday, M.A.K. (1967/8) 'Notes on Transitivity and Theme in English', *JL*, 3,4.

(forthcoming) 'Text as Semantic Choice in Social Contexts', in T.van Dijk and J.Petöfi eds, *Grammars and Descriptions.* Berlin.

Halliday, M.A.K. and R.Hasan (in press) *Cohesion in English.* London.

Hill, T. (1958) 'Institutional Linguistics', *Orbis*, 7.

Hymes, D. (ed. 1971) *Pidginization and Creolization of Languages.* Cambridge, C.U.P.

Jones, K. (n.d.) 'The Role of Discourse Analysis in Devising Undergraduate Reading Programmes in EST (English for Science and Technology)', CILT.

Killingley, S.-Y. (unpb) 'Internal Structure of the Cantonese Word and General Problems of Word Analysis in Chinese'.

Labov, W. (1970) 'The Study of Language in its Social Context', *Studium Generale*, 23.

Le Page, R.B. (1969) 'Problems of Description in Multilingual Communities', *TPS 1968.*

(1973) 'The Concept of Competence in a Creole / Contact Situation', *YPL*, 3:31-50.

(1975a) 'Sociolinguistics and the Problem of "Competence"', *Language Teaching Abstracts*, 8:137-56.

(1975b) 'Projection, Focussing and Diffusion, or, Steps Towards a Sociolinguistic Theory of Language', *Proceedings of the Conference on the Methodology of Sociolinguistic Surveys, Montreal.* Washington D.C.

Liu, J. (1962) *The Art of Chinese Poetry.* London.

Lyons, J. (1968) *Introduction to Theoretical Linguistics.* Cambridge, C.U.P.

Martin, J. (forthcoming) 'Phoricity: a Systemic Analysis of Reference'.

Pound, E. (1916) *Gaudier-Brzeska.* London.

Sacks, H. *et al.* (1973) 'A Simplest Systematics for the Organization of Turn-Taking for Conversation'. Mimeo.

Sankoff, G. (1972) 'Language Use in Multilingual Societies: Some Alternative Approaches', in J.Pride and J.Holmes eds, *Sociolinguistics: Selected Readings.* London, Penguin Books.

(1974) 'A Quantitative Paradigm for the Study of Communicative Competence', in R.Bauman and J.Sherzer eds, *Explorations in the Ethnography of Speaking.* Cambridge, C.U.P.

Shukman, A. (1975) 'Voloshinov and the Theory of Dialogue', *Neo-Formalist Circle 10th Colloquium* (abstract).

Sinclair, J. (*et al* 1972) *The English Used by Teachers and Pupils,* SSRC Project Report, Birmingham University.

Sinclair, J. and Coulthard, R.M. (1975) *Towards an Analysis of Discourse: the English Used by Teachers and Pupils.* Oxford U.P.

Ure, J.N. (1973) 'Register and Unit of Discourse', draft working paper, Language Use in Ghana (SSRC project). Legon.

(1974a) 'Code-switching and "Mixed Speech" in the Register-systems of Developing Languages', in *Proceedings of the Third International Congress of AILA, 1972, Copenhagen.*

(1974b) 'Self-reporting and the Question of Reliability', International Congress of Sociology, Toronto.

Wexler, P. (1964) 'On the Grammetrics of the Classical Alexandrine', *Cahiers de Lexicologie*, 4:61-72.

Widdowson, H. (1973) *An Applied Linguistic Approach to Discourse Analysis.* Edinburgh, unpbl. PhD thesis.

(1975) *Stylistics and the Teaching of Literature.* London.